

想定確信度の時系列的更新

松井 理直

神戸松蔭言語科学研究所・大阪保健医療大学

Author's E-mail Address: matsui@sils.shoin.ac.jp

Author's web site: <http://sils.shoin.ac.jp/~matsui/>

Time-series Updating of a Degree of Assumption

MATSUI Michinao F.

Shoin Institute for Linguistic Sciences, Osaka Health Science University

Abstract

認知主体が現実世界で情報獲得する際、情報を混乱させる外界のノイズや認知主体自身を持っている先入観など様々な妨害要因が存在する。本稿では、こうした妨害要因が存在する時に、認知主体が想定をどのように変更していく可能性があるかについて論じたものである。

When we acquire the right cognitive environment, we cannot search for every relevant piece of information because our cognitive ability is not limitless. Cognitive agents, however, must have proper cognitive environments from partial information structure. In order to achieve this purpose, we need to continue updating cognitive environments through our experience in real world. This paper proposes a valid algorithm for updating the environments on the basis of the estimated degree of belief, and discusses the qualitative feature of our innate cognitive ability.

キーワード：関連性理論、想定、認知環境の認知環境

Key Words: Relevance Theory, Degree of Assumption, Updating of the Cognitive Environment

1. はじめに

認知主体は現実の環境の中で、様々な情報を獲得していく。そうした情報のうち、いくつかの情報は真であり、知識としての価値を持つ。しかし、中には偽である情報も存在し、認知主体の知識体系を混乱させる原因となる。認知主体は様々な点で限界を持つ存在者であり、

全知全能の存在者でない故に、情報の真偽を完全に知ることはできない。しかし、そのような不完全な状況においても、認知主体は可能な限り体制化しようとする。そのためには、認知主体は経験と共に不適切な情報を排除し、適切な情報を強化する形で、認知環境を常に更新していかなければならない。松井（2010）では、こうした想定確信度に基づく認知環境のアルゴリズムについて議論を行った。このモデルは安定した情報獲得が可能な枠組みの一種であるが、仮説を反証するデータに比較的敏感であり、「誤った信念」を持ち続けることがない。しかし、固定観念や先入観に代表されるように、認知主体が「誤った信念」を持ち続けることがあることも確かであり、こうした「誤った」想定を変更しないことがあるという性質も、人間という認知主体の推論過程を考える上で重要な点の1つであろう。本稿では、松井（2010）の議論を改訂し、こうした時に過ちを犯す認知主体の情報獲得（すなわち認知環境の更新過程）に関する数理モデルについて議論を行う。

2. 誤った信念獲得について

2.1 関連性理論

まず議論における背景として、Sperber and Wilson（1986）による**関連性理論**（*Relevance Theory*）を採用する。この理論は次のような枠組みを持つ。まず、認知主体にとって、外界やあるいは自身の内部にある全ての情報が利用可能であるとは限らず、認知主体にとって**顕在的**（*manifest*）な情報のみが利用できる。こうした顕在的情報のことを特に**想定**（*assumption*）と呼ぶ。また、顕在的情報である想定の総体を**認知環境**（*cognitive environments*）という。認知主体は常に新しい情報を獲得し、また推論によって自主的に新しい想定を生み出す存在であるため、認知環境は常に変化し、更新されている。こうした認知環境の変化の中で、特に認知環境の「改善」をもたらす作用を**認知効果**（*cognitive effects*）と呼ぶ。認知効果は

- (1) a. 新しい想定の獲得
- b. 不確実な想定の確定
- c. 誤った想定の棄却

のいずれかによってもたらされる。この認知効果という点から、関連性理論の中心概念となる「関連性」について、「不必要なコストを払うことなしに認知効果をもたらすデータを、**関連性**（*relevance*）を持つ情報という」といった形で定義を行う。

認知主体にとって、この関連性という性質はとても重要なものである。人間という認知主体を取り巻く環境は極めて範囲が広く、かつ常に情報が流動的に変化している世界である。しかし、認知主体の持つ知覚・思考・伝達といった情報処理能力は、世界に存在する膨大な情報の一部分しか処理することができない。したがって、完全解を常に求めることができるとは限らない。そうした限界の中で、認知主体はを少しでもよりよい解を得るために、部分情報を手がかりにして可能な限り安定した体制化と推論を行う。認知の本質は、巨大な認知環境が内包する多様性に適応していく能力であり、無限の情報の中で、必要と思われる情報

を取捨選択する能力だといってもよい。

効率的に部分情報を処理するためには、当該の情報にとって無関係なものは無視し、関係のあるもののみを選択すればよい。Sperber and Wilson (1986) は、人間の認知系が自らにとって関連のある情報に注意を払うようデザインされていると見なし、認知主体のこうした性質を**関連性の認知原理** (*cognitive principle of relevance*) と呼んでいる。また、なお、情報の受け取りが動的に行われるコミュニケーションでは、関連性の伝達原理 (*communicative principle of relevance*, 情動的意図と伝達の意図) も重要な役割を果たす。

2.2 正しい想定としての知識

関連性理論における顕在的情報 (想定) は、「知識」であることが最も望ましい。一般に認識論では、「知識」すなわち「認識主体 S が命題 A を知っている」ためには、「A が真であるということ」を S が確信しており、A が実際に真であり、A が真ということ S が確信する当然の理由がある」ことが基本的条件であると考えられてきた (Ayer, 1981)。還元するなら、知識とは「正当化された真なる情報」ということになる。したがって、誤っていると分かっていることを信じ続けるのは知識ではない。また、自分では正しいと思いこんでいる情報であっても、それが実際に正しくなければやはり知識とはいえない。さらに、認知主体がある情報を真だと信じており、その内容が実際に真であったとしても、それが単なる「まぐれ当たり」であれば、やはり認知主体が知識を持っているとはいえない。信念や想定と知識との大きな違いはこの点にある。偽である何かを想定したり信じたりすることはできるが、偽である何かを「知る」ことはできない。

こうした知識の持つ「正当化」という条件については、多くの議論がなされてきた。例えば、正当化の条件を十分な証拠に基づくことと定義する「強い証拠主義」がその代表といえよう。しかし、強い証拠主義に立つと、ほぼ全ての想定が知識とは言えなくなる。帰納推論は不確実性を持ち、多くの証拠を集めても次の瞬間に反例が見つかるかもしれない。演繹推論でも、命題に十分な証拠を求め続けると、ある命題の証拠となる別の命題を次々と求めなければならず、直接感覚所与以外の情報について無限後退に陥る。直接感覚さえも、ゲティア問題 (Gettier, 1963) のように、常に正当化されるとは限らない。

この証拠主義の問題点を克服するために、正当化の条件を認識主体の活動のみに限定するのではなく、認識主体の外部に求めるという「外在主義」についても様々な議論がなされてきた。外在主義の代表的な考え方としては、信頼性主義や整合主義 (Armstrong, 1973) を挙げることができる。信頼性主義の正当化条件は「S の想定が信頼できるプロセスにより形成された場合」と定義される。整合主義では、「S の想定が他の想定と矛盾を起こさない」ことが正当化の条件となる。例えば、感覚などから直接得られる基本的情報は、それを歪める阻害要因が明確でない限り、信頼性主義により正当化された知識といってよい。また、ある想定が他の基本想定と矛盾しないのであれば、自分自身でその想定の本拠を説明できない場合でも、整合主義により正当化された情報であり、知識と見なし得る。

2.3 フランシス・ベーコンのイドラ論

しかし、情報獲得という動的な場面を考えた場合、信頼性主義や整合主義にも難しい問題が伴う。その主な理由は、情報獲得が演繹推論ではなく帰納推論によって行わざるを得ないこと、また認知主体が例外事象や偶発的な情報といったノイズが存在する現実世界の中で生きていることに起因する。イギリス経験主義の祖といわれるフランシス・ベーコン（1561～1626）は、主著『ノヴム・オルガヌム』の中で「イドラ論（idola）」を展開し、こうした認識上の錯誤について、以下の4種類があると指摘した。

(2) a. 種族のイドラ（自然性質に基づくイドラ）

錯誤の原因が人間という類（あるいは人間性）そのものにあるもので、人間の感覚における錯覚や人間の本性にもとづく偏見を指す。人類一般に共通して起こる錯誤であり、例としてミュラー・リヤーの錯視や明るさの恒常性などが挙げられる。

b. 洞窟のイドラ（個人経験に基づくイドラ）

各人に固有の本性や狭い経験によって、事象の解釈が歪められることに起因してする錯誤を指す。「カラスは黒い（実際には白いカラスが存在する）」といった思い込みや、「この鉛筆を使うと良いことがある」という迷信行動などが洞窟のイドラに相当する。

c. 市場のイドラ（伝聞によるイドラ）

人間同士のコミュニケーションから生じる錯誤で、噂やマスメディアが流す意図的な虚偽を信じてしまう例が相当する。

d. 劇場のイドラ（権威によるイドラ）権威者の発言や思想、伝統や通説等無批判に信じることから生じる偏見を指す。中世における天動説に基づく宇宙観や、ある種の宗教的信念などが相当する。

人間は、これらのイドラによって誤った思い込みを形成してしまうと、例えその思い込みに反する「正しい事例」が多く現れても、それらの事例を無視ないし軽視してしまう傾向を持つ。また、場合によっては、正しい情報を思い込みに合致するよう変形して解釈してしまう。これらは知識獲得において重大な障害であり、こうした錯誤を避けるために必要な帰納推論の妥当性について多くの議論がなされてきた。また、こうした錯誤を引き起こさないような知識獲得の手法についても、様々な数理モデルが提案されている。

しかし、人間がこうした錯誤を引き起こす存在である限り、関連性理論のような「人間」の認知理論を考える場合には、こうした認識上の誤りを引き起こすようなモデルが必要である。次節では、松井（2009）で議論を行った認知的関連性の計算モデルを拡張し、上記のイドラの中で特に認知的関連性と深い関係を持つ洞窟のイドラ・劇場のイドラについて考察してみたい。

3. 情報の真理判断における計算モデル

3.1 認知的関連性の最も単純な計算方法

ある発言 Y の意味を理解するということは、情報 Y と最も関連の深い文脈情報 X を同定する事に他ならない。これは（文脈）情報 X と（明示された）情報 Y との共起確率に基づいた次のような認知環境を最適化することに等しい。

表 1：認知環境における想定間の可能性

		情報 Y		
		Y	$\neg Y$	合計
情報 X	X	$P(xy)$	$P(x\bar{y})$	$P(x)$
	$\neg X$	$P(\bar{x}y)$	$P(\bar{x}\bar{y})$	$P(\bar{x})$
合計		$P(y)$	$P(\bar{y})$	1

松井（2009）では、こうした認知環境において、情報 X から情報 Y への最も単純な関連性計算は、次式で行えることを議論した。この式は統計学の観点からいえば、 $k_{\bar{x}} = 1$ の時に、情報 X から情報 Y への回帰係数を計算していることと等価であり、 $r = \sqrt{\beta(x \Rightarrow y) \cdot \beta(y \Rightarrow x)}$ によって情報 X と情報 Y の相関係数 r も求めることが可能である。また、Hattori（2003）で提案されている DH モデル（dual-factor heuristics model）における情報 X と情報 Y の因果関係を表す指標 H も、 $H(x \Rightarrow y) = \sqrt{P(y|x) \cdot P(x|y)}$ として計算できる。すなわち、 $\beta(x \Rightarrow y)$ は情報 X と情報 Y の関係を計算する最も基本的な式といってよい。

$$\begin{aligned}
 (3) \quad \beta(x \Rightarrow y) &= P(y|x) - k_{\bar{x}} \cdot P(y|\bar{x}) \\
 &= \frac{P(xy)}{P(x)} - k_{\bar{x}} \cdot \frac{P(\bar{x}y)}{P(\bar{x})} \\
 &= \frac{P(xy)}{P(xy)+P(x\bar{y})} - k_{\bar{x}} \cdot \frac{P(\bar{x}y)}{P(\bar{x}y)+P(\bar{x}\bar{y})}
 \end{aligned}$$

実際の意味理解・意味獲得においては、情報 Y は明示的に与えられているが、文脈情報 X はいくつかの候補が考えられ、決定されていない。したがって、 $P(x|y)$ および $P(\bar{x}|y)$ から、ベイズの定理に基づいて式（3）の計算を行うことになる。さらに情報 Y の確率は確定しているため、最終的に求めたい情報 X は、肯定情報と否定情報とのコントラストに基づいて決定されると見なしてよい。精神物理学における Weber-Fechner の法則や Stevens のベキ法則においても、情報のコントラストが心理量をもたらす基本情報となっていることから考えても、情報のコントラストに基づいて意味情報が決定されることは妥当であろう。

3.2 情報値・ランダムノイズ・偏向要因

次に、認知主体が探索対象として情報 X を選び、その情報値のコントラストに基づいて情報 X の想定確信度を決定する場合を考えてみよう。なお、この時に否定情報 $\neg X$ は探索対象とされない点に注意されたい。ここで、情報 X の情報値を変数 t で表す。 t は $0 \leq t \leq 1$ の範囲を取る実数で、情報 X が二値の命題である時には、 $t=1$ の時が真、 $t=0$ の時が偽を意味する。情報 X が三値の真理値を採る命題であるなら、 $t=0.5$ の時が未知 (*unknown*) に相当し、情報 X が多値の真理値や情報値を取る場合は、 t を $0 \leq t \leq 1$ の範囲で真理値 (情報値) に対応させる。また、外界における情報 X の生起を問題にしている場合には、 $t=1$ の時が実際に生起しており、 $t=0$ の時が生起していないことを示す。情報 X の生起が確率的に起こる場合には、 $t=P(x)$ と見なしてよい。

こうした情報 X の情報値 t に対し、情報 X の想定されるコントラスト T は $T = f(t) - (1 - f(t))$ 、すなわち $T = 2f(t) - 1$ で表せる。 $f(t)$ は $f(1) = 1, f(0) = 0, 0 \leq f(t) \leq 1$ ($0 \leq t \leq 1$ の条件に注意されたい) を満たし、かつ単調性を備えた関数であり、その結果 $-1 \leq T \leq 1$ を常に満たす。こうした関数 $f(x)$ を本稿では認識関数と呼んでおく。また、認知主体がコントラスト量 T に基づいて情報 X を受理する認知能力の鋭敏さを係数 a で表す。この係数 a を認識度指数と呼ぼう。認識度指数 a は経験値によって変動する実数で、良質な情報を経験するほど値が大きくなり、逆に悪質な情報を経験すればするほど値が小さくなる性質を持つ。

さて、理想世界ではなく、実世界での情報認識・情報獲得を考える場合には、コントラスト量 T に影響を与える係数として、認識度指数の他に情報 X を攪乱させる外部妨害要因の S/N 比を考えなければならない。これは、本稿において情報が二値の命題である時にも「真理値」(正確には情報値)として実数値 t を割り当てなければならない理由にもなっている。すなわち、認知主体の外で情報 X を攪乱させる外部妨害要因がある場合には、例え情報が真の二値命題であっても、その情報値は 1 にはならない。

さらに、こうした妨害要因を考える時に、その妨害が偶然発生するものか、あるいは認識過程において常に存在するものかという点も重要である。今、前者を外部妨害要因を「ランダムノイズ」、後者を「偏向要因」と呼ぼう。外界要因で情報 X と常に連動する情報は、ノイズではなく、何らかの相関関係や因果関係を持つ関連情報であると見なせる。したがって、観測対象情報 X に直接影響を与える阻害要因は、ランダムノイズと考えてよい。ここで、ランダムノイズの S/N 比を s とする時、認知主体が受容できる「有効な」情報 X の情報値は $(1 - \frac{1}{s}) \cdot t$ に等しい。この結果、ランダムノイズの中に存在する情報 X に対する認知主体の潜在的認識量 aT は、

$$(4) \text{ 潜在的認識量} : aT = 2a \cdot f\left(\frac{s-1}{s} \cdot t\right) - a$$

として計算できる。したがって、ランダムノイズがほとんど存在しなければ (すなわち S/N 比がほぼ無限大であれば)、認知主体は外界のコントラスト量に比例して情報を受容できるが、ランダムノイズの影響が非常に強ければ、情報の受容量はゼロに近づいてしまう。ただし、

情報 X の受容を妨害するランダムノイズと偏向要因のうち、ランダムノイズは外界が良質でなくても、情報 X のデータ数さえ集められれば無視することができると考えられる。

それに対して、偏向要因は情報 X を「常に」混乱させるため、情報認識や情報獲得に本質的な影響を及ぼす。偏向要因の代表例としては、先入観・迷信・誤った定説・支配権力やマスメディアによって意図的に流される虚偽情報などが挙げられるだろう。本稿では、こうした複数の偏向要因を区別せず、その有無をまとめを変数 N で表す。 $N=1$ は認識に影響を与える偏向要因が存在し、 $N=0$ なら偏向要因は関与しない。さらに、この偏向要因が認知主体に与える影響の大きさを実数値で表す偏向度指数 b を導入する。 b は実数値で、この数値が大きいかほど「偏向要因が真なる情報の受容を妨害する」程度が強いことを示す。この偏向度指数は、例えば誤った先入観が存在すれば $b > 0$ の値を取るが、たまたまその先入観が正しい場合には $b < 0$ の値を取ることに注意されたい。この結果、認識を誤らせる偏向量は単純に次式で表される。

$$(5) \text{ 偏向量} : = bN$$

3.3 想定確信度の計算式

ここで、認知主体が情報 X の真偽を判断できる可能性の程度を示す指標 V を考えよう。この判断可能性は認知主体が持ちうる認識量と偏向量の相互作用に依存して決まると仮定できる。したがって、判断の指標 V は

$$(6) V = aT - bN$$

として表現してよい。同時に、この指標 V は情報 X の想定確信度 $P_n(\bar{x})$ に基づくエントロピーによっても定義することができる。すなわち、情報 X に関する判断の容易さは、情報 X の想定確信度と否定情報 $\neg X$ の確信度がどの程度大きく違っているかに比例すると仮定してよい。したがって、最終的な認識における肯定情報と否定情報とのコントラスト量は、両者の相対的な比率（オッズ比）の情報量（ロジット）によって表される。このことから真理判断指標 V は、

$$(7) V = \log_2 \frac{P_n(x)}{P_n(\bar{x})}$$

に等しい。ここで、 V に関する 2 つの関係式 (6), (7) より、

$$(8) \log_2 \frac{P_n(x)}{P_n(\bar{x})} = aT - bN$$

すなわち

$$(9) \frac{P_n(x)}{P_n(\bar{x})} = \exp(aT - bN)$$

が成立する。この式 (9) は、情報 X の確信度と否定情報 $\neg X$ の確信度とのオッズ比 $\frac{P_n(x)}{P_n(\bar{x})}$ が、コントラスト量 T およびノイズ N と指数関数的な関係を持つことを意味している。さらに式 (9) に $P_n(\bar{x}) = 1 - P_n(x)$ を代入すると、式 (10) を得る。この式 (10) は式 (11) に等しい。

$$(10) P_n(x) = \frac{\exp(aT - bN)}{1 + \exp(aT - bN)}$$

$$(11) P_n(x) = \frac{1}{1 + \exp(-aT + bN)}$$

これを、もともとの情報 X の情報価 (真理値) t で表すなら、式 (12) となる。すなわち、情報価 (真理値) と想定確信度はロジスティックな関係を持つ。

$$(12) P_n(x) = \frac{1}{1 + \exp\left(a - 2a \cdot f\left(\frac{s-1}{s} \cdot t\right) + bN\right)}$$

4. 情報価と想定確信度の関係

4.1 認識度指数の影響

本節では、式 (12) を用いて、情報価と各種の妨害要因が想定確信度に与える影響について議論を行う。まず、認識度指数が想定確信度に与える影響について見てみる。の影響については、攪乱要因であるランダムノイズと偏向要因が一切ない状態を仮定するのが最も分かりやすい。式 (12) において、ランダムノイズの S/N 比 $s = \infty$ 、偏向要因 $N = 0$ の状態で、認識度指数 a を変化を様々に変化させてみた結果を、図 1 に示す。

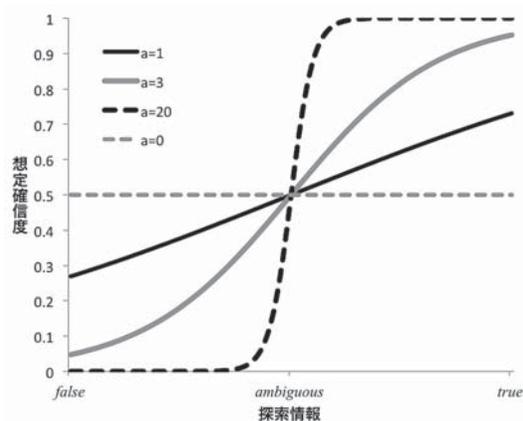


図 1：認識度指数が想定確信度に与える影響

いずれの場合も点対称のグラフになり、また情報価（真理値）が $t=0.5$ の時に、想定確信度も $P(x)=0.5$ となることが分かる。また、認識度指数が高いほど、“true”の情報に対する想定確信度はより 1 に近づき、“false”の情報に対する想定確信度はより 0 に近づく。すなわち、想定確信度は認知主体にとって多値の真理値として用いることができるような性質を持つ。

より詳しく認識度指数の影響を見てみよう。認識度指数が $a=1$ の場合は、真理値（あるいは情報価）と想定確信度が線形的に対応する。また、情報が真である時でも想定確信度は約 0.73、偽なる情報に対する想定確信度は約 0.27 となっており、想定確信度の変動範囲が狭い。このことは、どのような情報であっても想定確信度は比較的近似した値となり、情報の真偽判断が明確にできなくなることを意味している。なお言うまでもないが、認識度指数が $a=0$ の場合は、いかなる情報に対しても想定確信度は 0.5 となり、真偽の弁別は完全に不可能となる。

一方、認識度指数が強くなると、真なる命題に対しては想定確信度がほぼ 1 となり、偽なる命題に対する想定確信度はほぼ 0 に近づく。すなわち、真偽に関してほぼ完全な確信度が得られている。また、真理値が *unknown*（あるいは情報価が 0）に近い情報であるほど、ほんの少しの情報変化が想定確信度を大きく変動させ、真に近い情報あるいは偽に近い情報になるほど、情報変化が想定確信度に与える影響力が減少していることも特徴の一つである。この性質は、認識度指数が上昇するにしたがって、世界をなるべく安定した形で理解する能力も上昇することを示している。認識度指数が極めて大きくなると、 T が 0 以下なら想定確信度は 0 となり、 T が 0 以上なら想定確信度は 1 となる。すなわち、世界をほぼ完全に真か偽かに二分できるような認知環境が構築される。これは、いわゆる「範疇的知覚」に近い性質で、認識度指数の違いによって、クリスプなカテゴリーを構築するか、プロトタイプ・カテゴリーを構築するかの違いが生じることを示す。

こうした性質は、情報獲得の特性を考える上でも興味深い。例えば、認識度指数が小さい場合には、“true”の情報でも想定確信度は 1 にならず、“false”の情報に対しても想定確信度は 0 にならない。これは、経験が浅く、認識が十分に行われない場合には、「情報の価値を決めつけない」ことを意味しており、情報獲得の点で重要な性質である。経験が不十分な時に、「どちらかという真」「どちらかという偽」というこうした慎重な態度を取っておくことは、不必要な先入観や誤った常識の影響から免れるために欠くべからざる特質とってよいだろう。逆に、十分に経験を積んだ場合は、目前にあるたった 1 つの情報について、即座に真偽を決定することも、世界をよりよく認識するために大切なことである。低い認識度指数が慎重な想定確信度を与え、高い認識度指数が範疇的な想定確信度を与えるという傾向は、経験と共に正しく効率のよい情報獲得を行う上で、必要不可欠な性質であるといつてよいだろう。

4.2 S/N 比の影響

次に、情報 X と同じ外界に存在するランダムノイズの影響を見てみよう。この影響の強さは S/N 比を表す係数 s の強さによって示され、 s が小さな値であればあるほど、妨害要因としてのランダムノイズのラウドネスが強い。この S/N 比の影響を図 2 に示す。

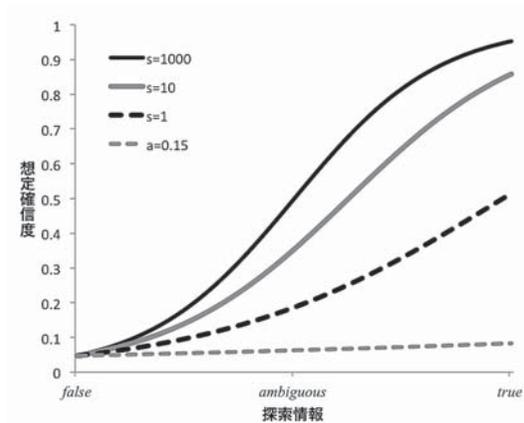


図2：S/N比の影響

S/N比が小さくなるにつれて、すなわちライダムノイズの影響が強くなるにつれて、真理値が“true”である情報が認知主体の内部では「偽」として判断されるようになっていくことが分かる。ノイズが強くなるほど、探索情報を見つけにくくなるため、これは自明のことと言えよう。これに対し、“false”の情報は、S/N比の強さがどのようなものであれ、認知度指数が同一である限り、認知主体の内部では同一の強さで「偽」と判断されている。この“false”の情報に対し、ランダムノイズが何の影響も与えないという性質は、認知主体の外部にあるランダムノイズと、認知主体の内部に起因する先入観や固定観念との違いを考える上で、非常に重要な点である。図2から分かる通り、S/N比が極端に低い場合には、現実の情報値・真理値がどのようなものであれ、認識上は全て「ほぼ偽である」と判断されてしまう。これは、どのような情報でもダメだと思いつく一種の先入観のような性質とも理解できよう。しかし、非常に微妙な点で、強いランダムノイズの影響は先入観の影響とは異なる性質を持つ。それが、この“false”の情報に対する影響に表れるのである。この点については、4.4節で再度議論を行う。

4.3 認識関数の影響

式(12)において、認識関数だけは他のパラメータとは全く異なった性質を持つ。認識度指数であれS/N比であれ、次節で述べる偏向度指数であれ、これらは想定確信度に一定の影響を与える係数に過ぎない。すなわち、外界を認知主体の内部に取り込む時に局所的な影響を与えるに過ぎない。しかし、認識関数は認知主体の内部世界のあり方そのものを決定する要因である。

実際の認識関数がどのような関数系であるのかは、現時点では明確ではない。しかし、参考になる先行研究として、外界情報を心理情報に変換する精神物理学的関数としてよく知られている Fechner の法則と Stevens のベキ法則を挙げることができるだろう。前者は、物理世界を対数軸に従う心理世界に変換する関数である。比率尺度に従う等比級数的な変化を持つ

物理世界を、感覚尺度で構成される等差級数的な変化を持つ心理世界に変換する関数といってもよい。グラフで言うと、Y軸の片対数グラフで表示される特性を持つ。一方、Stevensのベキ法則は、グラフでいうと両対数グラフで表示される性質を示す。すなわち、比率尺度に従う物理世界を比率尺度で構成される心理尺度に変換する関数系である。本稿で議論している想定確信度は、0が想定しないことを示す値となっているため、比率尺度に近い性質を持つ。したがって、本稿ではベキ法則に近い形で認識関数の特性を設定することにしよう。Stevens型のベキ関数が、認識関数の性質すなわち単調性を持ち、かつ $f(1)=1, f(0)=0, 0 \leq f(t) \leq 1$ を満たすという特性を備えている点も偶然ではない。ここで、認識関数におけるベキ指数の影響を図3に示す。ベキ法則と同じく、ベキ指数に1を持つ認識関数を用いた場合には、 $t=0.5, P(x)=0.5$ のポイントで点対称のグラフとなり、真偽判断が公平に行われることが分かる。これに対し、ベキ指数が1より小さい認識関数が適用された場合には、認知主体は「真」に偏想定確信度を持つ。すなわち、「偽」であることがかなり明確である場合にのみ「偽」に近い判断が下される。一方、ベキ指数が1より大きくなればなるほど、「真」と判断される世界は縮小され、情報価が0.5を大きく上回っても依然として「偽」と判断する傾向が続く。言い換えるなら、認識関数におけるベキ指数が1より小さい場合には「疑わしきは罰せず」という原則に基づいた判断となり、ベキ指数が1より大きくなればなるほど疑い深い判断を下す傾向を持つことになる。

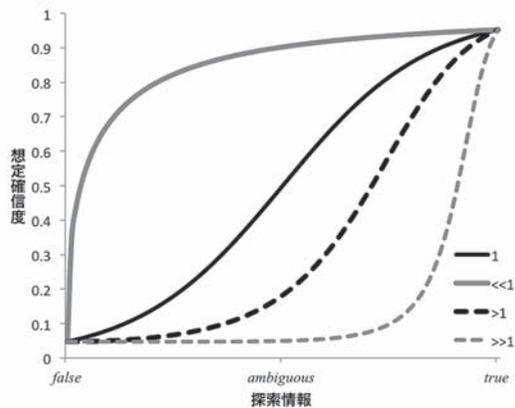


図3：認識関数の影響

4.4 偏向要因の影響

式(12)において最後に残るパラメータは、偏向要因に関する bN である。前述したように、この偏向要因はランダムノイズと同じく、情報価や真理値の正しい認識に混乱をもたらす。しかし、その影響がランダムノイズとは異なる点に注意が必要である。偏向度指数が想定確信度に与える影響を図4に示す。偏向度指数 b が0より小さくなるほど、どのような情報価であっても「真」という想定判断を下す可能性が高くなり、逆に b が0より大きくなるほど、

どのような情報に対しても「偽」という判断を下す可能性が高くなっていることが分かる。

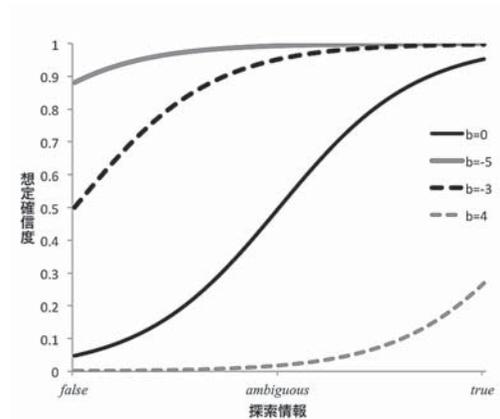


図4：偏向要因の影響

既に述べた通り、想定確信度がアンバランスになる傾向は、ランダムノイズのS/N比が小さい場合や認識関数のベキ指数が1からずれた時にも生じやすい。しかし、偏向度指数が想定確信度に与える影響は、他のパラメータが及ぼす影響とは全く異なる性質を持つ。図5は、認識度指数を $a=2$ に固定した時に、偏向度指数、S/N比、ベキ指数が単独で各々どのような影響を判断に与えているのかを示したものである。

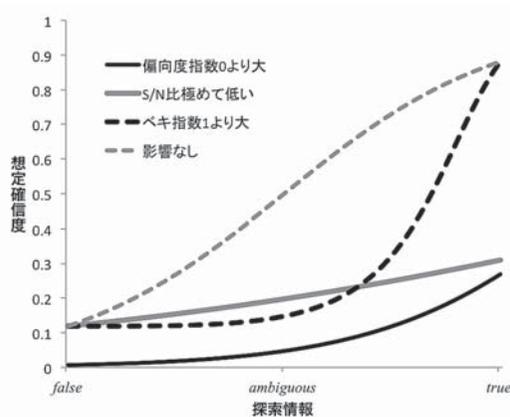


図5：各種パラメータが偽に偏る判断に与える影響

偏向度指数が0より大きい場合とS/N比が非常に小さい場合、ベキ指数が1よりも小さい場合のいずれにおいても、これらの影響がない時（図5の薄い点線）に比べて「偽」に偏った判断が行われているのが分かる。しかし、S/N比およびベキ指数の作用は、情報価が0で

ある時や真理値が完全に“false”である時には全く何の影響を与えていない。さらにベキ指数の場合には、情報価が1である時や真理値が完全に“true”である時にも、攪乱要因の影響がない場合の想定確信度と等しい値を持つ。これに対し、偏向度指数が0より大きい場合には、 $t=0$ の場合においてもさらに確信度を下げる効果を与えていることが分かる。もちろん、他の情報価であっても、基準となる判断よりも常に確信度が低い。

この性質は、式(12)における偏向要因が「拭いがたい先入観」や「強固な偏見」「固定観念」といった概念に相当するものであることをよく示している。すなわち偏向度指数は、外界の情報そのものよりも認知主体の内部に存在する情報に強く依存した判断を行う要因の影響を示すものといってよい。これに対し、1より小さいベキ指数は「よい疑い深い判断を下す」だけであり、外界の情報が完全であれば（すなわち完全に偽であったり完全に真であったりすれば）、通常の想定と同一の判断となる。一方、S/N比は認知主体の内面に起因する問題ではなく、外界におけるノイズの影響が強すぎて、そもそも外界情報を受容できないことによって判断の狂いが生じているに過ぎない。

こうした性質は、「positiveな判断を下しやすい」場合を考えるとさらに分かりやすい。図6は、偏向度指数がマイナス値の場合、S/N比が極めて大きい場合、認識関数のベキ指数が1より大きな場合の結果を示す。図5とは異なり、まずS/N比（ランダムノイズ）は標準的な想定確信度（図6の薄い点線）に限りなく近づくが、それを「上回る」ことがない。すなわち、ランダムノイズは「よりpositiveな判断を下しやすい」方向には影響を与えない。これは、ランダムノイズが「情報をマスクする」効果しか及ぼさないことに起因する。こうしたS/N比の影響に対し、偏向要因は「どんな情報でも真と判断しやすい」傾向も生み出してしまう。ランダムノイズも偏向要因もいずれも判断を攪乱する妨害要因ではあるが、外部の妨害要因であるランダムノイズと認知主体内部の妨害要素である偏向要因とは本質的に異なる性質を持つのである。なお、認識関数のベキ指数の影響については、ベキ指数が1より小さい場合と同じく、外界の完全情報（完全な真あるいは完全な偽）に対しては通常の想定と同一の判断を行う。前節で述べたように、ベキ指数が1より小さい時の認識関数における真偽判断の偏りは、「疑わしきは罰せず」という判断傾向を意味しており、S/N比と同じく、先入観や固定観念とは異なる性質を持つのである。

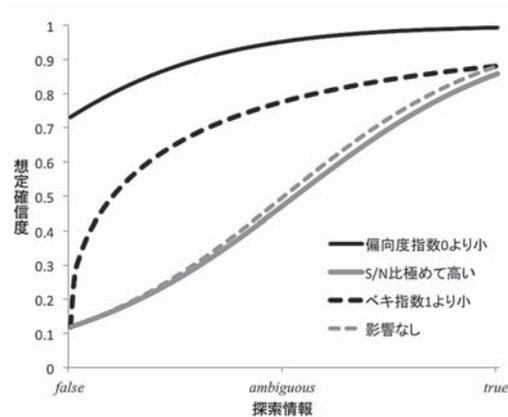


図6：真に偏る判断における各種パラメータの影響

5. 想定確信度の時系列的更新

5.1 認知環境の更新

認知主体は時間の流れの中に存在しており、そこで様々な経験を踏まえながら正しい情報を獲得していく。こうした過去・現在・未来という時間における情報のうち、過去に生じた情報については無知であっても未知ではない。少なくとも過去の各情報は真か偽かのいずれかであることは確信できるので、過去の世界に対して適当な信念確信度を割り振ることができる。また、今現在直接経験している情報については、それを完全に真なるものと即決することは難しいにせよ、かなり信頼性の高い情報と見なしてよいだろう。これに対し、未来の情報は既定のものとは異なるものがあり、後者については過去・現在から予測するしかない。こうした時間の流れの中で、ある情報が「一般的」に（すなわちトークンとしてではなく、タイプとして）どの程度真理であり得るかを判断する場合、例えば次のようなシステムが考えられる。

- (13) a. 初めて経験する情報に関して、暫定的に一つの世界像を構築する。
- b. 現在直接経験した情報 X に関する想定確信度は、過去に経験した情報 X の想定確信度を元に、より「真」である方向に重み付けした想定確信度を設定する。
- c. 未来の情報に関しては、過去と現在の情報から暫定的な予測を行う。
- d. これらの想定確信度を総合した形で、現時点での想定確信度を決定し、経験を経る度に刻々と更新していく。

具体的な変更アルゴリズムは以下のようなになる。

- (14) a. 現実世界における n 回目の証拠が情報 X であった場合、情報 X に関する確信度は以下の通りである。

1. 過去 ($n-1$ 回目まで) の累積情報の確信度: $P_{n-1}(x)$
 2. 現在の確信度: $P_n(x)$
 3. 近接未来に対する確信度: $P_{n-1}(x)^2$
- b. 同様の場合、情報 $\neg X$ に関する確信度は以下の通りである。
1. 過去の累積情報に対する確信度: $P_{n-1}(\bar{x}) = 1 - P_{n-1}(x)$
 2. 現在の確信度: 0
 3. 近接未来に対する確信度: $1 - P_{n-1}(x)^2$
- c. 現実世界における n 回目の証拠が情報 $\neg X$ であった場合、情報 X に関する確信度は以下の通りである。
1. 過去の累積情報に対する確信度: $P_{n-1}(x)$
 2. 現在の確信度: 0
 3. 近接未来に対する確信度: $1 - P_{n-1}(\bar{x})^2$
- d. 同様の場合、情報 $\neg X$ に関する確信度は以下の通りである。
1. 過去の累積情報に対する確信度: $P_{n-1}(\bar{x}) = 1 - P_{n-1}(x)$
 2. 現在の確信度: 1
 3. 近接未来に対する確信度: $P_{n-1}(\bar{x})^2$

また、式 (12) によって想定確信度を更新していく場合、認識度指数・認識関数のベキ指数・変動度指数のいずれか (複数でもよい) のパラメータ値を変更することになる。なお、言うまでもなく、ランダムノイズの S/N 比については、認知主体の外部で起こることなので、能動的なパラメータ値の変更は行われぬ。次節以降で、各々 100 試行のシミュレートを行い、それぞれのパラメータが時系列上の獲得過程に与える影響について考察を行う。条件を揃えるために、出現する情報の真値は二値 (すなわち $t=1$ か $t=0$) に限定し、このうち、以下の時点で「偽」の情報、すなわち想定に対する反例が出現した条件を設定する。

- (15) a. 全試行回数は 100 回。
- b. 11 回目と 21 回目に反例が出現。
 - c. 31 回目～35 回目まで 5 回連続して反例が出現。
 - d. 81 回目～85 回目までも 5 回連続して反例が出現。

5.2 認識度指数のみを変更していく場合

まず、経験によって認識度指数のみを変更した場合の獲得過程を図 7 に示す。グラフから分かる通り、情報 X に関する想定確信度は経験回数と共に上昇しており、真の情報を 9 回連続して経験した時点で確信度は 0.8 を超える。また、12 回目～20 回目、22 回目～30 回目を見て分かる通り、事前想定 of 確信度がより高い状況で真の情報を 9 回連続して発見すると、さらに高い想定確信度に到達し、認知主体は当該情報がこの世界においてはほぼ真であるという想定を持つ。

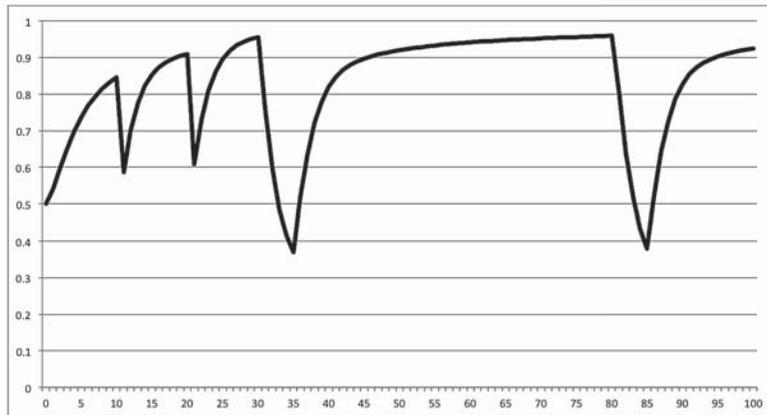


図7：認識度指数に依存する獲得過程

一方、反例の影響については、グラフから、たった1回でも反例であっても、その時点で想定確信度が0.6程度の値にまで下がる。これは想定確信度が0.9を越えている場合（21回試行目の反例）においても同様であり、反例が連続する場合には想定確信度が0.4程度まで下降する。このことは、式（12）に基づく想定確信度が単なる生起確率ではなく、反例となる否定的情報により敏感に反応することを示す。また、反例に出会うタイミングが遅い場合であっても、この性質は大きく変わらない。これは、仮説に反する例はその仮説の誤りを積極的に示唆するものであるに対し、仮説に反しない例は状況証拠に過ぎないという自然科学の在り方から見ても、妥当な計算といってよいであろう。

5.3 認識関数におけるベキ指数のみを更新する場合

次に、経験値が認識関数のベキ指数に直接影響を与える場合の結果について、図8に結果を示す。基本的な獲得傾向は図7に類似していることが見て取れる。ただし、認識度指数の影響と異なる点として、まず情報獲得のスピードが速いことが上げられよう。10回連続して正事例を経験した場合、認識度指数に基づく獲得では想定確信度は0.8を越える程度であるが、ベキ指数を変化させる獲得ではほぼ0.9に到達している。

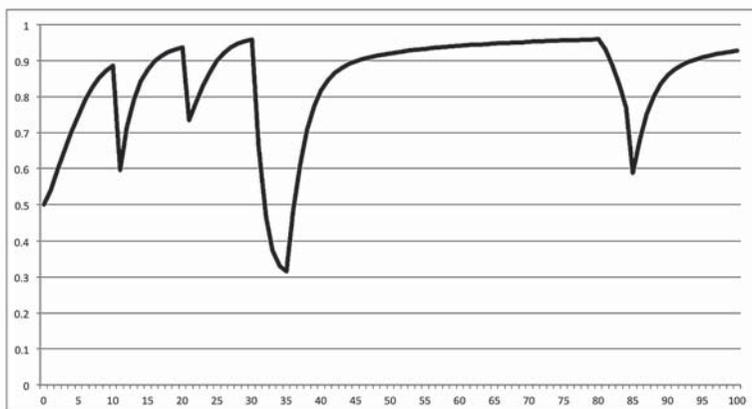


図 8：認識関数のベキ指数に依存する獲得過程

認識度指数と認識関数のベキ指数が持つ効果の違いが最も顕著に見られるのは、反例の影響があった時である。図 7 と図 8 における第 21 試行目の反例および第 81 回目～85 回目の反例部分を見ていただきたい。反例を経験した時点で、想定確信度が下がっている点は共通するが、認識度関数のベキ指数に基づく想定獲得では、認識度指数の効果ほど劇的な影響はもたらさない。このことは、認識度指数が現前の情報をより重視する性質を持つのにに対し、認識関数のほうが過去の累積経験により依存した獲得を行うことを示している。このことは、反例の連続回数が少ない場合を考えると、さらに分かりやすい。図 9 は 81 回目の試行から 5 回連続ではなく 3 回連続して反例が出現した場合のシミュレーション結果であるが、認識関数のベキ指数に依存した情報獲得では、この程度の反例に対しては仮説を保持する傾向が強い。

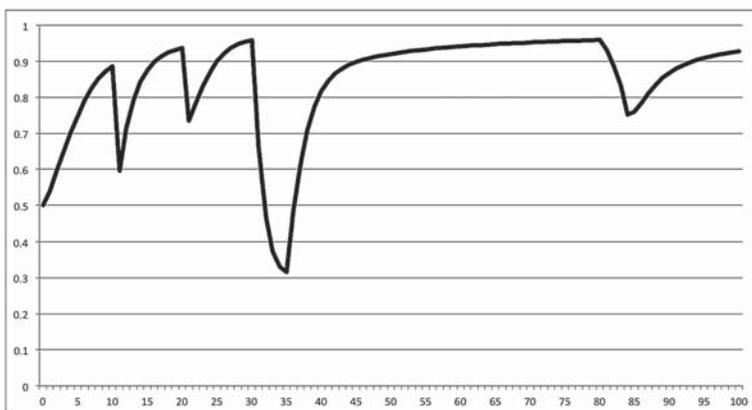


図 9：81 回目からの反例連続回数が少ない場合

科学哲学の観点から考えると、反例は仮説に対して「絶対的」な影響を持つ。還元するなら、客観的な判断を行うためには、認識関数のベキ指数を1から大きく動かしてはならない。しかし、化学実験ほど環境が統制されていない現実の場面では、偶然に何らかの妨害を受ける可能性を否定できない。そのため、一見反例と思われる現象が存在しても、現在保持している仮説を即座に変更するのは必ずしも合理的な態度とは言えず、現時点での仮説を基本的に保持しておくほうが妥当ということも考えられるであろう。認識関数の型はこうした比較的保守的な判断に近い効果をもたらすものと考えられる。

5.4 偏向要因のみを更新する場合

最後に、過去の経験が偏向要因にのみ影響を与える獲得過程について見てみよう。シミュレーションの結果を図10に示す。反例の与える影響が、これまでの結果と大きく違うことが見て取れる。

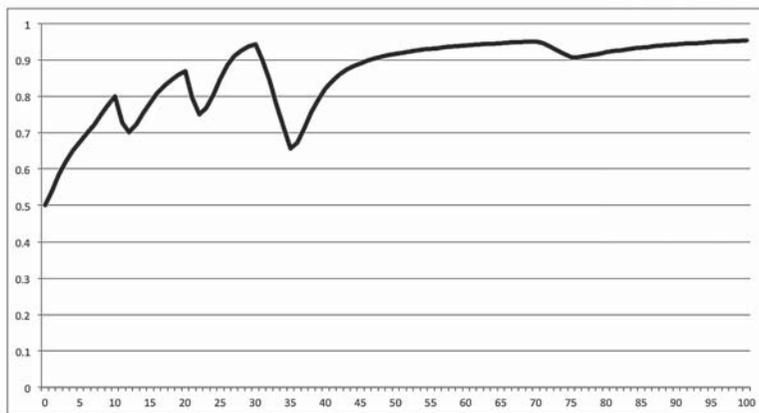


図10：81回目からの反例連続回数が少ない場合

特に十分な正事例を経験した81回目からの反例については、ほとんど何の影響も与えていない。想定に対する確信度が0.9を越えるような経験が十数回以上続いた場合、その想定がほぼ定着してしまい、その後に連続した反例を経験したとしても、(誤っている可能性のある)想定を保持し続けてしまう。これが、偏向要因が一種の先入観や固定観念に相当する効果であるということの意味である。

偏向要因のみがこうした影響を与える理由は、図5や図6から明白であろう。すなわち、認識度指数や認識関数のベキ指数(あるいは外界ノイズのS/N比)は、様々な認知バイアスを与えはするが、外界情報が完全である時($t=1$ か $t=0$ の時)には、その情報にほぼ対応した想定確信度を生み出す。しかし、偏向度指数のみは、この指数が大きくなると、外界情報がどのようなものであれ、常に認知主体の内部に存在する想定にこだわってしまうのである。

6. 総合論議

本稿では、人間の認識錯誤に注目しながら、妥当な想定確信度の計算モデルについて検討を行った。その結果、外界情報の情報価あるいは真理値に対してロジスティックな形で想定確信度を決定するモデルを用いると、種々の錯誤をシミュレートできることが確かめられた。用いたパラメータは以下の通りである。

- (16) a. 情報価に直接影響する S/N 比：外界に存在する攪乱要因（ランダムノイズ）の強さを表す。
- b. 認識関数のベキ指数：S/N 比の影響を受けた情報価の判断価値について、線形・非線形に拡大縮小を行う。
- c. 認識度指数：認識関数の出力結果に掛かる係数を意味する。
- d. 偏向度指数：認知主体の内部に存在する攪乱要因（固定観念や先入観など）の強さを表す。

これらのいずれもが、何らかの形で外界の情報を「ゆがめてしまう」可能性を持つ。例えば、(16a) は認知主体の外部において認識形態とは独立に存在するランダムノイズであり、認知主体にとって制御することが困難な要素である。ただし、化学実験のように認知主体が観察する外部環境自体を理想化できる状況下では、こうした要因を除去できるだろう。あるいは、ノイズがランダムに変化する要因であるので、何度もデータを取得することでノイズの影響を相殺したり、統計的に S/N 比の見当もつけられるようになる。S/N 比が分かれば、正しい想定値を逆算できるようになり、したがって、フランシス・ベーコンのイドラに関する議論 (2) では、この要因は考慮されていない。

これに対し、(16b)、(16c)、(16d) は認知主体内部に存在する要因であり、「イドラ」の原因になり得る。例えば、「種族のイドラ (2a)」の一種である「知覚過程における錯覚現象」は精神物理学的関数によってもたらされるものであるため、(16b) における評価関数の形式そのものに起因するイドラと言ってよいだろう。また、「洞窟のイドラ (2b)」は真偽に関して非対称な判断に起因する錯誤であるが、経験によって修正のしやすい錯誤であり、これは (16b)、(16c) によって引き起こされるものと考えられる。これに対し、「劇場のイドラ (2d)」はいわゆる偏見や通説、先入観に相当する錯誤であり、これまで議論してきたように (16d) の作用によって起こるものと見てよい。すなわち、これらのイドラは関連性理論における

「関連性の認知原理」が関与する錯誤と言える。なお本稿では議論を行わなかったが、「市場のイドラ (2c)」はコミュニケーションに起因する錯誤であり、関連性の伝達原理がより強く関わるエラーと言えよう。

さらに、式 (12) に基づく想定確信度計算が正しいとすると、「種族のイドラ」「洞窟のイドラ」「劇場のイドラ」のうち、前者2つは世界が完全であれば（すなわち真理値が1か0であることが明確であれば）、認識の歪みをもたらさない。これは図1および図3において、

false と true の想定確信度が常に安定していることから言えることである。しかし、「劇場のイドラ」は世界が完全であったとしても、外部世界の情報に依存した判断が行われるとは限らない。そういう意味では、先入観や固定観念という錯誤は、一旦形成されてしまうと拭いがたい認知の歪みであり、そうした性質は式 (12) によってうまく表現できる (図 5 および図 6 を参照のこと)。

以上の議論から、式 (12) は認知環境を構成するための想定確信度を計算する枠組みとして、ある程度妥当な性質を持っているとよいであろう。この式に従う時系列上の獲得においても、(16b), (16c) のパラメータを変化させている限り、初期の経験時点では想定確信度が 0.5 前後に設定され、時間と共に想定確信度が 1 や 0 に近づいていく。また、仮説の反証例に出会った時には、頻度確率に関わらず、想定確信度を一旦 0.5 近くのリセットできる。こうした性質は、経験から正しい想定を見つけ出す上で有利な特質とよい。もちろん、(16d) のパラメータを変化させた時には、時系列上の獲得過程であっても、外界情報を正しく判断することができなくなっていく。すなわち、偏見などに引きずられた判断を行ってしまうことになる。しかし、こうした性質も認知主体が時に持つてしまう判断の歪みであり、我々の認知特性をシミュレート上でも考慮しなければならない性質であろう。

なお、本稿では論じなかった認識錯誤の興味深い特性として、式 (12) における認識関数 $f(x)$ の特性そのものに起因する歪みを挙げるができる。また、「市場のイドラ」はコミュニケーションの特性を考える上で避けて通れない認識錯誤の性質であろう。これ点につ謝い辞では、関連性の伝達原理に関する形式化と共に、また稿を改めて議論してみたい。

謝辞

本研究は、日本学術振興会科学研究費補助金・基盤研究 (C) 「認知的関連性のモデル化と文理解実験に基づく実証的研究」(平成 22 年度～平成 25 年度、研究代表者：松井理直、課題番号：22520415)、による援助を受けている。

文献

- Armstrong, David M. (1973). *Belief, Truth and Knowledge*. Cambridge University Press, Chicago.
- Ayer, A.J. (1981). 『知識の哲学』. 白水社.
- Gettier, Edmund L. (1963). Is Justified True Belief Knowledge?. *Analysis*, **23**, 121-123.
- Hattori, Masashi (2003). Adaptive Heuristics of Covariation Detection: A Model of Causal Induction. *Proceedings of the 4th International Conference on Cognitive Science and the 7th Australasian Society for Cognitive Science Joint Conference (ICCS/ASCS 2003)*, **1**, 163-168.
- 松井理直 (2005). 計算論的関連性理論における日本語条件文の解釈. *Theoretical and Applied Linguistics at Kobe Shoin*, **8**, 95-122.

松井理直 (2007). 計算論的関連性理論に基づく日常的推論の分析. *Theoretical and Applied Linguistics at Kobe Shoin*, **10**, 45-76.

松井理直 (2009). 認知的関連性の単純かつ妥当な計算方法. *Theoretical and Applied Linguistics at Kobe Shoin*, **12**, 21-36.

松井理直 (2010). 認知環境の更新に関する妥当な計算手法について. *Theoretical and Applied Linguistics at Kobe Shoin*, **13**, 23-49.

Sperber, Dan & Wilson, Deirdre (1986). *Relevance: Communication and Cognition*. Blackwell.

内田聖二ほか訳 (1993). 『関連性理論—伝達と認知—』. 研究社出版.

(受付日 : 2014. 1. 10)